Model-based real-time optimisation of a fed-batch cyanobacterial hydrogen production process using economic model predictive control strategy

Ehecatl Antonio del Rio-Chanona, Dongda Zhang*, Vassilios S. Vassiliadis

Department of Chemical Engineering and Biotechnology, University of Cambridge, Pembroke Street, Cambridge CB2 3RA, UK.

*: corresponding author, email: dz268@cam.ac.uk, tel: 44 (0)1223 330132.

Abstract

Hydrogen produced by microorganisms has been considered as a potential solution for sustainable hydrogen production for the future. In the current study, an advanced real-time optimisation methodology is developed to maximise the productivity of a 21-day fed-batch cyanobacterial hydrogen production process, which to the best of our knowledge has not been addressed before. This methodology consists of an economic model predictive control formulation used to predict the future experimental performance and identify the future optimal control actions, and a finite-data window least-squares procedure to re-estimate model parameter values of the on-going process and ensure the high accuracy of the dynamic model. To explore the efficiency of the current optimisation methodology, effects of its essential factors including control position, prediction horizon length, estimation window length, model synchronising frequency, terminal region and terminal cost on hydrogen production have been analysed. Finally, by implementing the proposed optimisation strategy

into the current computational fed-batch experiment, a significant increase of 28.7% on hydrogen productivity is achieved compared to the previous study.

Keywords: biohydrogen production, economic model predictive control, finite-data window least-squares, on-line optimisation, dynamic simulation, fed-batch process.

1. Introduction

Hydrogen is considered as one of the fuels with great potential to provide clean energy for transport, electricity and heating in the future (Tamburic et al. 2011). At present, microorganisms such as green algae, cyanobacteria and purple non-sulphur bacteria have been extensively studied for biohydrogen production (D. Zhang et al. 2015; Basak & Das 2006). Many efforts have been conducted to identify biohydrogen synthesis metabolic mechanisms in different species (Melis et al. 2000; Min & Sherman 2010; Bandyopadhyay et al. 2010). Effects of culture composition and light intensity on biomass growth and hydrogen production have also been comprehensively explored from both simulation and experiment aspects (Basak & Das 2006; Dechatiwongse et al. 2014; Oh 2004; D. Zhang et al. 2015; Dongda Zhang, Pongsathorn Dechatiwongse, et al. 2015; Tamburic et al. 2012b) to determine the favourable conditions for biogas production. In addition, to facilitate the industrialisation of biohydrogen production process, a variety of novel photobioreactors (PBR) with different configurations have been designed to enhance hydrogen production and biomass density (Wang et al. 2013; Tamburic et al. 2011; Basak et al. 2014).

In order to accomplish the scale-up of hydrogen production from laboratory to industry, long term biogas production process has been conducted in recent studies. For example, a 21-day and a 23-day fed-batch process for green algal hydrogen production have been reported by (Vijayaraghavan et al. 2009) and (Kim et al. 2010), respectively. A 31-day continuous process for cyanobacterial hydrogen production has been carried out by (Dechatiwongse et al. 2015). Similarly, a 24-day and a 30-day fed-batch process have also been developed by (Lee et al.

2011) and (Boran et al. 2010), respectively. Based on these studies, it is found that biogas productivity differs from 1.8 mL g^{-1} (biomass) h^{-1} to 37.9 mL g^{-1} (biomass) h^{-1} depending on both species nature and process operating conditions.

Despite these achievements, the low biogas productivity shown in recent studies still presents an open challenge for the industrialisation of biohydrogen production, and to fill this gap process optimisation becomes an indispensable tool to maximise the process performance. However, as hydrogen is only generated by green algae and cyanobacteria in anaerobic and nutrient-deprived cultures whilst cell growth happens in aerobic and nutrient-sufficient environments (Dechatiwongse et al. 2015), the incompatibility between biomass growth and biogas production conditions significantly complicates the optimisation of this process. As a result, although recent studies have tried to extend cell growth period and increase hydrogen production (Tamburic et al. 2012a; Tamburic et al. 2013), the results suggest that it is difficult to accurately estimate and control the addition of limiting nutrients during the entire process purely based on experiments. Therefore, to address this open challenge, a real-time dynamic optimisation framework has to be implemented.

Dynamic optimisation is the procedure of finding the optimal control by a given performance index (*e.g.* objective function) for a time-varying process. It has been extensively used for a number of off-line tasks in bioprocess simulation, including estimating parameter values for fermentation kinetic models (Dongda Zhang, Del-Rio Chanona, et al. 2015; Adesanya et al. 2014), identifying desired operating conditions for batch and fed-batch processes (Del

Ródenas-Chanona et al. 2015; Alagesan et al. 2013), conducting operating studies in response to disturbances and upsets, and exploring the design of control systems (Biegler 2014).

In spite of the wide application on off-line optimisation, it is notable that bioprocesses in general are networks of complex biochemical reactions manipulated by enzymes and affected by culture conditions, in which advanced regulation methods have to be carried out to ensure the performance and efficiency of the process. As a result, traditional off-line control may not be suitable for the optimisation of complicated bioprocess since small deviations between the on-going process and the expected behaviours can lead to significant losses in terms of process efficiency (Mailleret et al. 2004).

Model predictive control (MPC), on the other hand, is an on-line control implementation which is by now a well-established method for the optimal control of linear and non-linear systems (NMPC) (Grüne & Pannek 2011). This method has become the most widespread advanced control methodology currently used in industry (Anon 2013). The method approximates the solution of an infinite horizon optimal control problem, which is computationally intractable in general, by a sequence of finite horizon optimal control problems where the dynamic behaviour of the system is optimised over a prediction horizon by computing the optimal inputs over a control horizon (shown in Fig. 1). Then the first element of the resulting control sequence is implemented in each time step to generate a closed-loop static state feedback (Grüne & Pannek 2009).

In particular, the Economic MPC (EMPC) approach is that in which an economic criterion (*e.g.* profitability, efficiency, production, etc.) is directly included in the performance index of the MPC formulation. This implementation achieves higher accuracy for process optimisation compared to the conventional MPC method, since it systematically determines the optimal operating strategy based on the real time economic measurements whilst accounting for state constraints, input constraints and time-varying constraints (Biegler 1998; Ellis & Christofides 2013; Bemporad & Morari 1999). Therefore, in the current study an EMPC strategy is employed to maximise biohydrogen production in a -real-time framework by administering the optimal influent nutrient flow rate over the entire process.

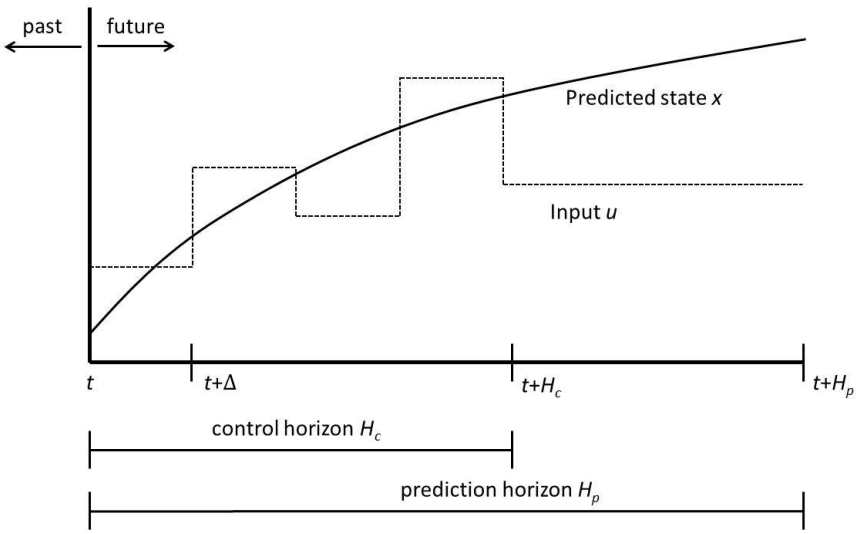


Figure 1: Model predictive control framework

2. Methodology Theory

2.1 Dynamic model for cyanobacterial hydrogen production

In the current study, cyanobacterium *Cyanothece* sp. ATCC 51142 is selected due to its high hydrogen productivity (Bandyopadhyay et al. 2010). In our previous study (Dongda Zhang,

Pongsathcrn Dechatiwongse, et al. 2015), a dynamic model has been proposed to simulate the entire batch process from cyanobacterial photo-heterotrophic growth to hydrogen production, and is shown in Equations (1a) to (1m). The detailed experimental setup and model construction can be found in (Dongda Zhang, Pongsathcrn Dechatiwongse, et al. 2015). Parameter values in the model are listed in Table 1. When simulating a fixed volume fed-batch process where dense nitrate (0.5 mol L^{-1}) and glycerol (0.1 mol L^{-1}) solutions are fed into the reactor, Equations (1b) and (1f) are replaced by Equations (1n) and (1o).

$$\frac{dX}{dt} = \bar{k}(I) \cdot \mu_{max,h} \cdot \left(1 - \frac{k_q}{q}\right) \cdot X \cdot \frac{C}{C + K_C} - \mu_{d,h} \cdot X^2 \quad (1a)$$

$$\frac{dN}{dt} = -Y_{N/X} \cdot \bar{k}(I) \cdot \mu_{max,h} \cdot \frac{N}{N + K_N} \cdot X \quad (1b)$$

$$\frac{dq}{dt} = Y_{q/X} \cdot \bar{k}(I) \cdot \mu_{max,h} \cdot \frac{N}{N + K_N} - \bar{k}(I) \cdot \mu_{max,h} \cdot \left(1 - \frac{k_q}{q}\right) \cdot q \cdot \frac{C}{C + K_C} \quad (1c)$$

$$\frac{dH_2}{dt} = Y_{H/X} \cdot \bar{h}(I) \cdot X \cdot f(N) \cdot f(O) \quad (1d)$$

$$\frac{dO_2}{dt} = Y_{O/X} \cdot \bar{k}(I) \cdot \mu_{max,h} \cdot \frac{N}{N + K_N} \cdot X - Y_{Od} \cdot \mu_{d,h} \cdot X^2 \cdot (1 - f(O)) \quad (1e)$$

$$\frac{dC}{dt} = -Y_{C/X} \cdot \bar{k}(I) \cdot \mu_{max,h} \cdot \left(1 - \frac{k_q}{q}\right) \cdot X \cdot \frac{C}{C + K_C} - Y_C \cdot X \quad (1f)$$

$$k(I) = \frac{I}{I + k_s + \frac{I^2}{k_i}} \quad (1g)$$

$$\bar{k}(I) = \frac{1}{L} \cdot \sum_{i=1}^{19} \left(k(I_0) + 2 \cdot k(I_i) + k(I_{20})\right) \quad (1h)$$

$$h(I) = \frac{I}{I + k_{s,H2} + \frac{I^2}{k_{i,H2}}} \quad (1i)$$

$$I = I_0 \cdot \exp\left[-\left(\tau_c \cdot X + \frac{3 \cdot \alpha_g}{d_b}\right) \cdot z\right] \quad (1j)$$

$$\bar{h}(I) = \frac{1}{L} \cdot \sum_{i=1}^{19} \left(h(I_0) + 2 \cdot h(I_i) + h(I_{20})\right) \quad (1k)$$

$$f(N) = 0.5 \cdot \frac{((N - 100)^2)^{0.5} - (N - 100)}{((N - 100)^2 + 0.1)^{0.5}} \quad (1l)$$

$$f(O_2) = 1 - \frac{O_2}{\left({O_2}^2 + 0.1\right)^{0.5}} \quad (1m)$$

$$\frac{dN}{dt} = -Y_{\frac{N}{X}} \cdot \bar{k}(I) \cdot \mu_{max,h} \cdot \frac{N}{N + K_N} \cdot X + F_N \cdot N_{in} \quad (1n)$$

$$\frac{dC}{dt} = -Y_{C/X} \cdot \bar{k}(I) \cdot \mu_{max,h} \cdot \left(1 - \frac{k_q}{q}\right) \cdot X \cdot \frac{C}{C + K_C} - Y_C \cdot X + F_C \cdot C_{in} \quad (1o)$$

where $\mu_{max,h}$ and $\mu_{d,h}$ are the maximum specific photo-heterotrophic growth and decay rate, respectively. q is the normalised nitrogen quota. k_q denotes the normalised minimum nitrogen quota. N, C, H_2, and O_2 are the concentration of nitrate, glycerol, hydrogen and oxygen, respectively. K_N and K_C are the half-velocity constant of nitrate and glycerol, respectively. $Y_{q/X}$, $Y_{N/X}$, $Y_{C/X}$, $Y_{O/X}$. $Y_{H/X}$, Y_C, Y_{Od} are the yields of nutrients and products as indicated by the respective subscripts of the symbols used. k_s and k_i are the light saturation and photoinhibition terms, respectively, for biomass growth. $k_{s,H2}$ and $k_{i,H2}$ are the light saturation and photoinhibition terms, respectively, for hydrogen production. τ_c is the light absorption coefficient by cyanobacteria, α_g is bubble volume fraction, d_b is average bubble diameter, I_0 is incident light intensity and I is the local light intensity cells experience in the reactor. F_N and F_C are nitrate and glycerol feeding rate (mL h^{-1}), respectively; N_{in} and C_{in} are influent nitrate and glycerol concentration, respectively. z is the distance after the exposure area, and L is the photobioreactor width.

The current species can only generate hydrogen after the depletion of oxygen and nitrogen source, when the culture condition switches from aerobic to anaerobic. Meanwhile, a sharp change in the activity of different *Cyanothece* sp. metabolic pathways is induced during the culture environment transition (Bandyopadhyay et al. 2010). To ensure biomass maintenance, nitrate has to be added during the process, which then leads to the generation of oxygen. As a

result, the culture will shift between the two conditions in a long term biogas production process (Feng et al. 2010). In addition, it is necessary for glycerol to be in excess in the culture since it provides the electrons for hydrogen production, and its concentration should be maintained between 20 mM to 50 mM (Min & Sherman 2010).

Since all of these factors may cause great deviations between the on-going experimental performance and model based prediction results, synchronising parameter values with the ongoing real process (parameter re-estimation) turns out to be the most effective and straightforward way to ensure the accuracy of the dynamic model predictions. The detailed introduction of parameter re-estimation approach is presented in Section 2.3.

Table 1: Parameters in the hydrogen production model.

Parameter	Value	Parameter	Value
$\mu_{max,h}$ h^{-1}	0.332	K_N mg	50.0
$\mu_{d,h}$ $L \cdot h^{-1} \cdot g^{-1}$	0.00716	$Y_{N/X}$ $mg \cdot g^{-1}$	492.7
k_q	0.165	$Y_{q/X}$ g^{-1}	0.0317
$k_{s,H2}$ $\mu mol \cdot m^{-2} \cdot s^{-1}$	140	$Y_{H/X}$ $mL \cdot g^{-1} \cdot h^{-1}$	14.20
$k_{i,H2}$ $\mu mol \cdot m^{-2} \cdot s^{-1}$	457	$Y_{O/X}$ $L \cdot g^{-1}$	81.02
Y_{Od} $L \cdot g^{-2}$	486.03	$Y_{C/X}$ $mmol \cdot g^{-1}$	20.454
Y_C $mmol \cdot g^{-1} \cdot h^{-1}$	0.0301	K_C $mmol \cdot L^{-1}$	0.0
α_g	0.0067	d_b m	0.002
k_s $\mu mol \cdot m^{-2} \cdot s^{-1}$	165	τ_c $m^2 \cdot g^{-1}$	0.126
k_i $\mu mol \cdot m^{-2} \cdot s^{-1}$	457	I_0 $\mu mol \cdot m^{-2} \cdot s^{-1}$	92.0

L m	0.025	T ℃	35.0

2.2 Computational experiments design

In the current study, a 21-day fed-batch computational experiment is designed to develop the real-time optimisation framework. The initial biomass concentration and nitrate concentration are fixed at 0.2 g L^{-1} and 150 mg L^{-1}, respectively. Incident light intensity and temperature are assumed to be 92 μmol·m^{-2}·s^{-1} and 35℃, respectively, which are the same with those in our previous research (Dongda Zhang, Pongsathcrn Dechatiwongse, et al. 2015). The current computational biohydrogen production process is then optimised by an Economic MPC strategy such that the controller dictates the optimal nitrate and glycerol inflow rates at every day.

The dynamic system has been discretised through orthogonal collocation over finite elements in time in order to solve the optimal control problems involved (Kameswaran & Biegler 2008). This results in a high-order implicit discretisation method which provides accurate profiles with relatively few finite elements, as well as good stability when dealing with stiff differential-algebraic equation (DAE) systems which encapsulate the process dynamics. Through orthogonal collocation the optimal control problems are transformed into large-scale nonlinear programming problems (NLP), and an interior point optimiser for large-scale nonlinear programs (IPOPT) developed by (Wächter & Biegler 2005) is used to solve the optimal solution in the present work.

Model results based on the parameter values listed in Table 1 are assumed to be the computational experiment performance of the 21-day hydrogen production process. A 10% error is embedded in the original parameter values to generate new parameter values for the dynamic model simulation at the beginning of the experiment. It is also assumed that state measurement errors are present, following a normal distribution with mean zero and standard deviation of 5% the measured value. Three computational experiments are carried out in the present study, and the optimised hydrogen production is then estimated by averaging the results of the three sets of computational experiments.

2.3 Model identification

To reliably optimise hydrogen production in real time it is important to have a robust model identification procedure which is able to estimate parameters efficiently at every stage of the process. Weighted non-linear least-squares method is one of the most efficient algorithms to estimate parameter values in a real-time optimisation framework where the data presents no gross errors (Zhang & Chen 2015). When these errors might exist in the measurements other methods such as redescending estimators (Arora & Biegler 2001), errors-in-variables-measured formulation (EVM) (Zavala & Biegler 2006), or other robust M-estimators can be employed to guarantee the accuracy of parameter estimation procedure. Given the high accuracy of current measurement instruments and the assumption that this measurement noise follows a normal distribution with mean zero and standard deviation of 5% the measured value, a weighted non-linear least squares algorithm is formulated to re-estimate the parameters.

Furthermore, due to the necessity of parameter re-estimation throughout the entire process, the computational effort will consistently grow with time as more information (experimental results) are collected and used in the non-linear least-squares procedure. To prevent the continuous increase in computational effort, a least-squares over a finite-data estimation window (FDW-LS) is employed in the present study. Once the optimisation problem is discretised and has been adapted to the orthogonal collocation formulation, the resulting NLP is defined as follows (Equations (2a) to (2h)).

$$\min_{p} \sum_{i=t-w+1}^{N} \left(\hat{x}_i - x(t_i, p)\right)^T \alpha_i \left(\hat{x}_i - x(t_i, p)\right) + \Delta p^T c_{\Delta p} \quad (2a)$$

subject to:

process dynamics

$$\dot{x}_{i,j} = f\left(x_{i,j}, \dot{x}_{i,j}, p\right) \quad (2b)$$

collocation constraints

$$x_{i,j} = x_{i-1,K} + h_i \sum_{l=1}^{K} \varphi_l(\tau_j)\, \dot{x}_{i,l} \quad (2c)$$

continuity constraints

$$x_{i,0} = x_{i-1,K} \quad (2d)$$

initial conditions

$$x_{1,0}(t_0) = x_0 \quad (2e)$$

integration horizon

$$0 \le t \le t_f \quad (2f)$$

bounds

$$x_{lb} \leq x \leq x_{ub} \quad (2g)$$

$$p_{lb} \leq p \leq p_{ub} \quad (2h)$$

where x is the vector of variables containing the chemical species in the model given by Equation (1a) to (1j), $\hat{x}$ is the measured states, p is the vector of the model parameters to be determined, w is the data estimation window length, α is the weighting factor, Δp is the change in parameters from the past iteration and $c_{\Delta p}$ is the penalty from prior parameter deviation. This optimisation problem is solved every time a new control input is supplied to the system, and old data points are discarded whenever they exceed the estimation window length.

2.4 EMPC on-line optimisation

EMPC is an on-line control strategy based on numerical optimisation. Future control actions and future model responses are predicted using a system model, and optimised at regular intervals with respect to a performance index. Given the FDW-LS strategy, future behaviour of the process can be predicted reliably, and the plan for the best actions can be identified accurately through EMPC. In the case of biohydrogen production, an optimisation problem is formulated such that at each day the hydrogen production at the end of the prediction horizon is maximised. This process is executed iteratively until the last day of the process. For the EMPC implementation in this work, a constraint such that hydrogen production must be higher than 100 mL L^{-1} at the end of the prediction horizon has been added as a terminal region constraint. In addition, the use of a terminal cost (quadratic objective function shown

in Equation (3b)) has been explored. The detailed formulation of terminal region and cost and their performance are presented in Section 3.5. The EMPC problem formulated for this work is the following (Equations (3a) to (3e)):

$$\max_{F_N(t),F_C(t)} H(t_{EMPC}) \quad (3a)$$

or

$$\min_{F_N(t),F_C(t)} (H(t_{EMPC}) - 2000)^2 \quad (3b)$$

subject to:

process dynamics

Equations (1a) to (1j)

bounds

$$0 \leq F_N(t) \leq 0.5 \text{ mL h}^{-1} \quad (3c)$$

$$0 \leq F_C(t) \leq 0.5 \text{ mL h}^{-1} \quad (3d)$$

terminal region

$$H(t_{EMPC}) > 100 \text{ mL L}^{-1} \quad (3e)$$

where t_{EMPC} is the final time in the EMPC prediction horizon. Once the sequence of optimal inputs from the optimisation problem are obtained only the first input is implemented as a control action, and the process is iteratively repeated for the whole duration of the experiment. The real-time optimisation strategy used in this work is depicted in Fig. 2.

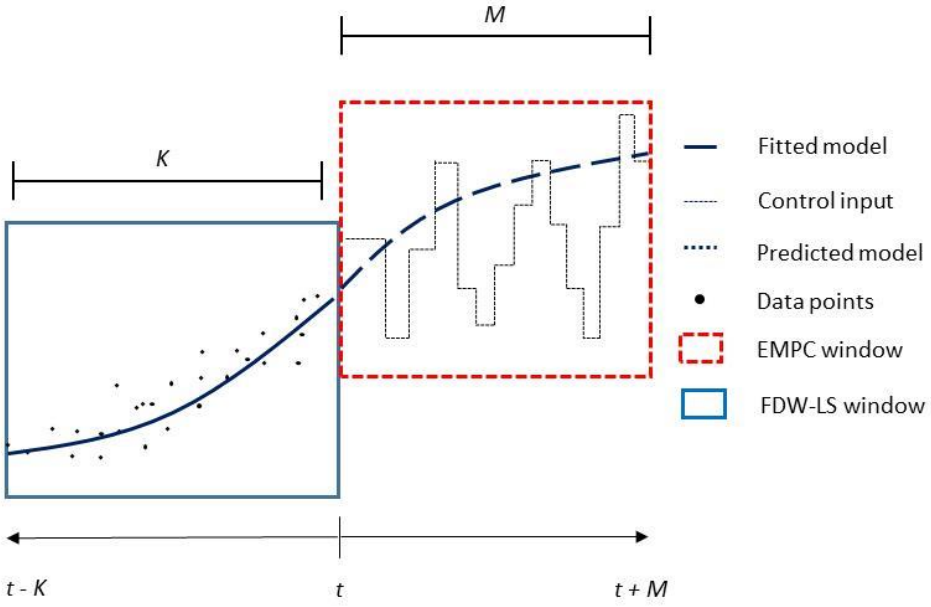


Figure 2: Schematic representation of the on-line optimisation framework

3. Results and discussion

3.1 Results of on-line optimisation

Fig. 3 shows the on-line optimisation result of the computational experiment with an estimation window length of 3 days and a prediction horizon of 5 days. Fig. 4 shows the inflow rates of nitrate and glycerol during the computational experiment. Both nitrate and glycerol inflow rates are changed once per day. Compared to the previous batch process under the same light intensity (Dechatiwongse et al. 2015), the current on-line optimisation experiment shows a 28.7% increase on hydrogen productivity.

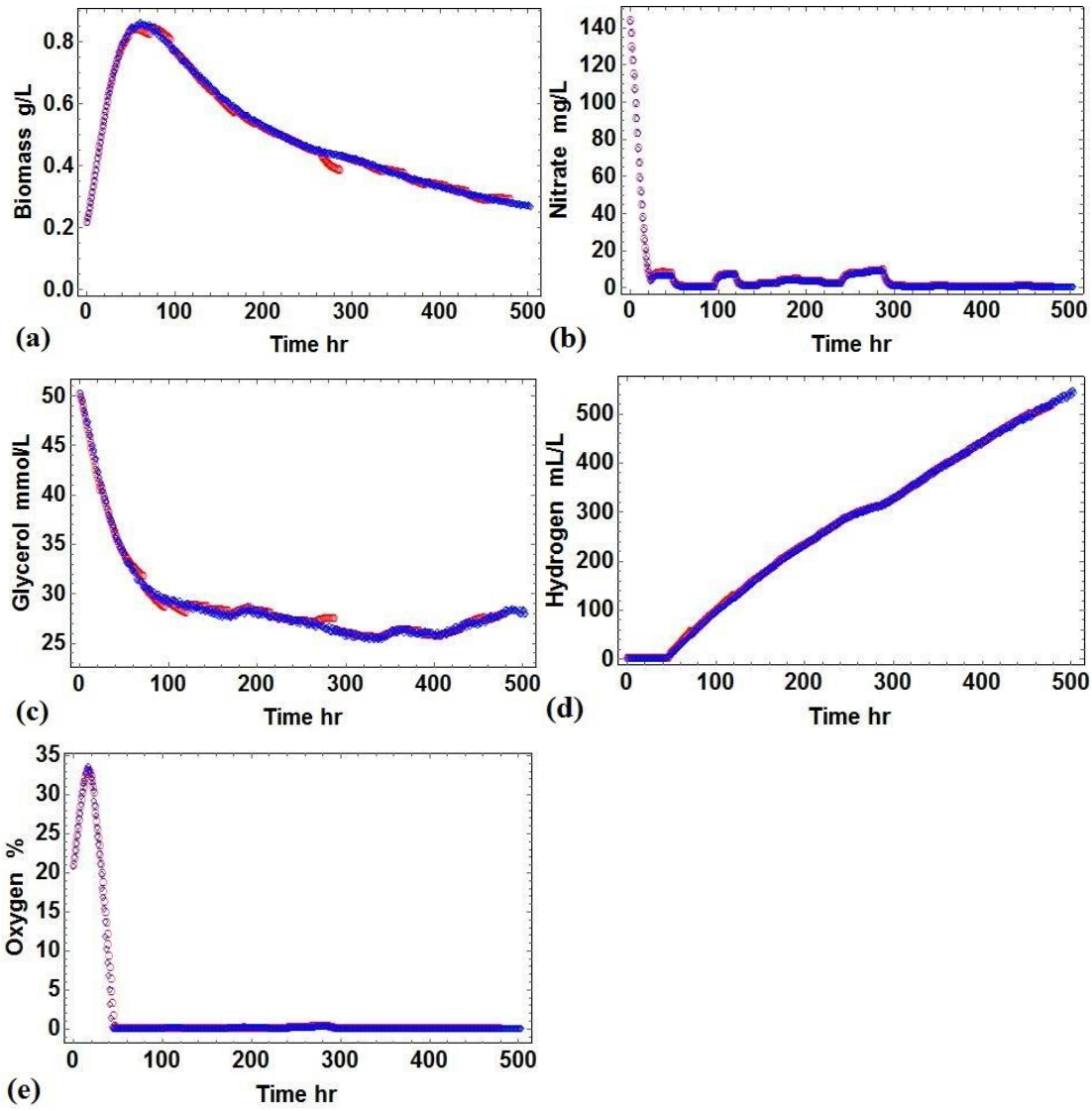


Figure 3: On-line optimisation result of the computational experiment. Estimation window length: 3 days, prediction horizon: 5 days, inflow rates: changed once per day. (a): biomass concentration; (b): nitrate concentration; (c): glycerol concentration; (d): hydrogen concentration; (e): oxygen concentration. Red points: model prediction; blue points: computational experimental result.

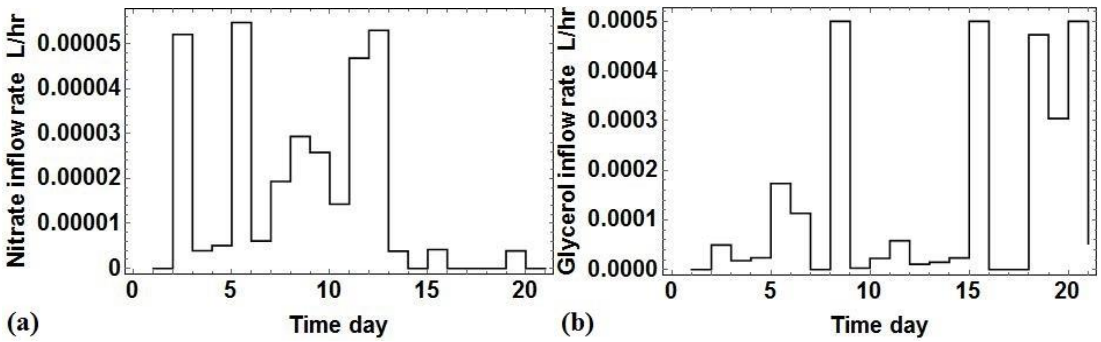


Figure 4: Inflow rate of nitrate and glycerol. Estimation window length: 3 days, prediction horizon: 5 days, inflow rates: change once per day. (a): nitrate inflow rate; (b): glycerol inflow rate.

From Fig. 3, it can be seen that the FDW-LS strategy is able to guarantee the dynamic model predicting accurately the experiment performance in the next 24 hours throughout the whole process. During the on-line optimisation process, all of the input nitrate is consumed by biomass for their maintenance, and nitrate concentration (Fig. 3b) in the reactor is kept low to ensure the culture is anaerobic (Fig. 3e) for continuous hydrogen production (Fig. 3d). Glycerol concentration in the reactor is also kept around 25 mM so that it is always in excess (Fig. 3c). Biomass concentration (Fig. 3a), however, decreases after the 100^{th} hour due to the limiting concentration of nitrate. Both nitrate and glycerol inflow rates (Fig. 4a and 4b) are precisely controlled during the entire time course of the operation so that total volume increase at the end of the process is negligible.

3.2 Effect of control position and prediction horizon

Fig. 5 shows the comparison of on-line optimisation experiments with different control positions. In one implementation the 5 controls were allocated to the first five days in the

prediction horizon length, and the other implementation distributes the 5 controls in equal time intervals (for example, for a 10-day prediction horizon experiment, the controls are allocated at the beginning of day 1, day 3, day 5, day 7 and day 9). The comparison therefore aims to identify the influence of control positions on optimal biogas production.

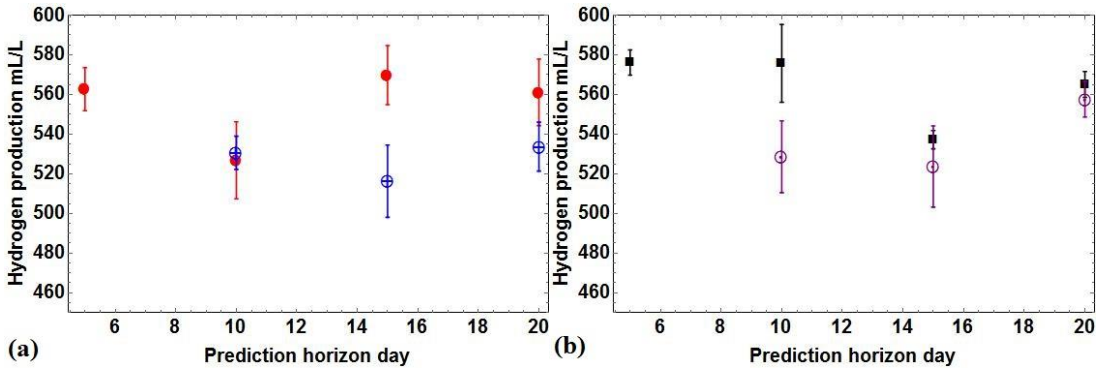
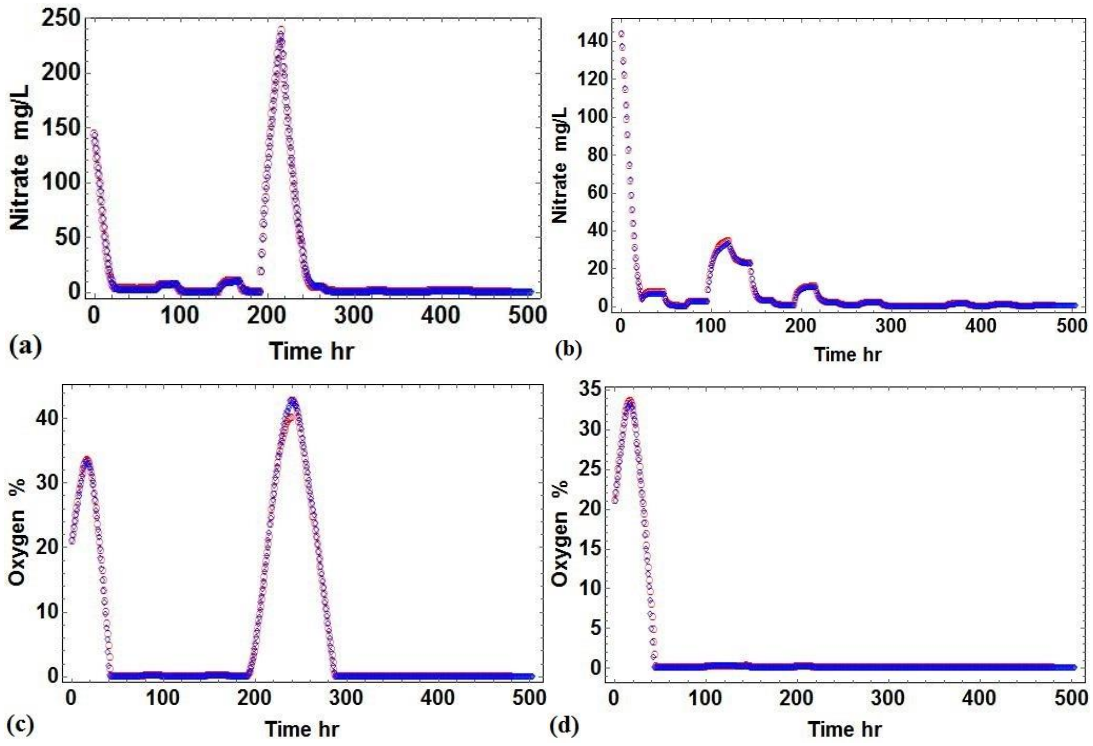


Figure 5: Comparison of on-line optimisation result with different control positions, both computational experiments have 5 controls. (a): experiments have an estimation window length of 3 days, filled circle: equal control interval, cross circle: controls only on the first five days; (b): experiments have an estimation window length of 5 days, filled square: equal control interval, open circle: controls only on the first five days.

From Fig. 5, it can be seen that experiments in which controls are equally spaced in the prediction horizon show higher hydrogen production than those whose controls are allocated in a daily fashion only in early days, although this increase is not very significant (less than 10%). This shows that controls can be placed more sparsely to increase the control horizon, without compromising process efficiency or controllability.

For example, Fig. 6 compares the performance of two computational experiments which

controls are assigned at different positions. From the figure, it can be found that when all the controls are allocated within the first five days in a 20-day prediction horizon, there is a dramatic increase on nitrate concentration after the 200th hour (Fig. 6a). This consequently leads to a sharp increase on oxygen concentration (Fig. 6c). As a result, hydrogen production is inhibited during a short period (Fig. 6e, around 80 hours) and its total production is decreased. On the contrary, when the controls are assigned with equal spacing, both nitrate concentration (Fig. 6b) and oxygen concentration (Fig. 6d) are kept low so that hydrogen can be continuously produced (Fig. 6e). Therefore, using equally spaced controls throughout a longer control horizon will be considered as the optimisation strategy for future work.



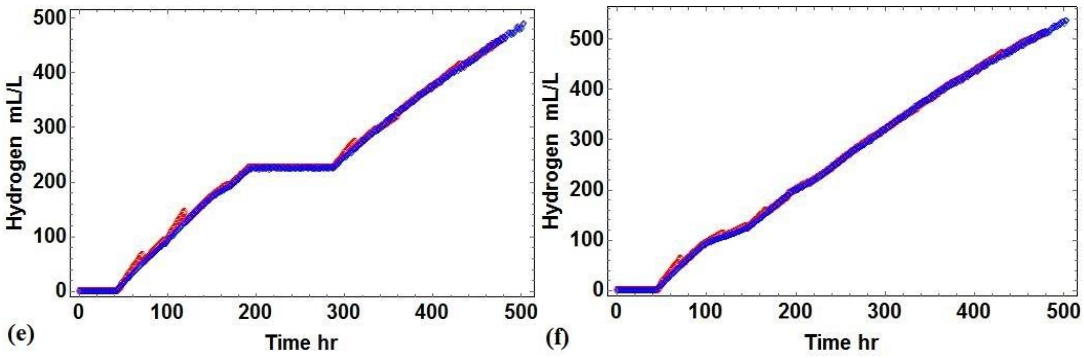


Figure 6: Comparison of computational experiment performance with different control positions. Both experiments have an estimation window of 3 days, a prediction horizon of 20 days and 5 controls. Red points: model prediction; blue points: computational experimental result. (a), (c) and (e): biomass concentration, nitrate concentration and hydrogen production when controls are in first five days; (b), (d) and (f): biomass concentration, nitrate concentration and hydrogen production when controls are equally allocated in the prediction horizon.

In addition, from Fig. 5a and 5b it is also found that the optimal biogas production is not significantly affected by the chosen prediction horizons if controls are allocated with equal space. As a result, a 5-day prediction horizon will be selected for future work due to its convenience.

3.3 Effect of estimation window length

Fig. 7 shows the comparison of optimisation result when the computational experiment estimation window length is 3 days and 5 days, respectively. From Fig. 7a, it can be concluded that changing estimation window length does not have a significant effect on the optimised hydrogen production, since two of the optimal biogas production are higher in the

5-day estimation window experiment whilst others are higher in the 3-day estimation window experiment.

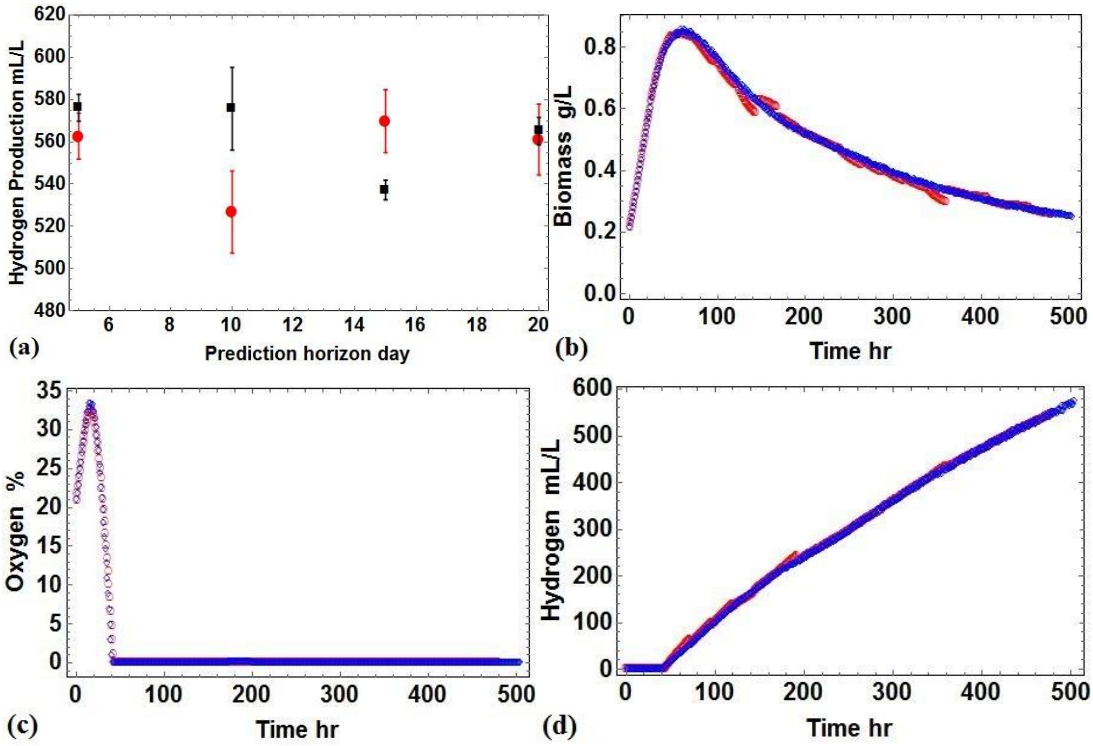


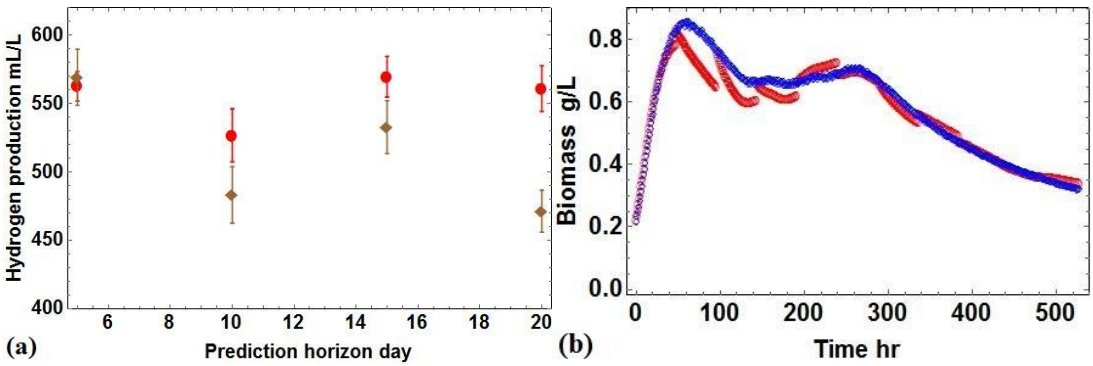
Figure 7: On-line optimisation result of the computational experiment with a 5-day estimation window and 5 controls. (a): comparison of the optimal hydrogen production with different estimation window length. Five controls are equally spaced in each horizon. Filled square: 5-day estimation window; filled circle: 3-day estimation window. (b), (c) and (d): biomass concentration, oxygen concentration and hydrogen production in the experiment with a 5-day estimation window and prediction horizon. Red points: model prediction; blue points: computational experimental result.

As the estimation window length can be directly linked to the accuracy of the dynamic model used for future process predication and optimal control estimation, the current comparison

suggests that a 3-day estimation window is long enough to guarantee the high accuracy of present model for on-line optimisation. Indeed, by comparing the model accuracy in the 5-day estimation window experiment (shown in Fig. 7b, 7c and 7d) with that in the 3-day estimation window experiment (shown in Fig. 3a, 3e and 3d), it can be concluded that both models are capable of well representing and predicting the experiment performance, since the errors between experiment and simulation results in both cases are negligible. Furthermore, the dynamic performance in both experiments are also similar. Therefore, the estimation window length in the future work will be selected as 3 days as it is less computationally demanding.

3.4 Effect of model modification frequency

Fig. 8 shows the comparison of optimisation result with different model parameter update frequencies. From Fig. 8a, it can be clearly concluded that optimal hydrogen productions from the experiment renewing model parameter values once per day are much higher than those from the experiment where model parameter values are only re-estimated once per two days.



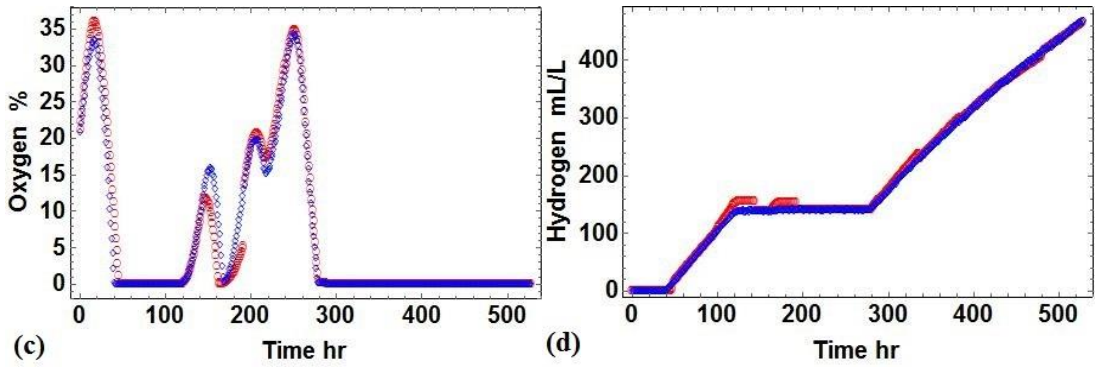


Figure 8: Optimisation result of the computational experiment in which model parameter values are re-estimated once per two days. The five controls are assigned with equal space and the estimation window is 3 days. (a): comparison of optimisation result with different parameter re-estimation frequency. Filled circle: update once per day, filled diamond: update once per two days. (b), (c) and (d): biomass concentration, oxygen concentration and hydrogen production during the time course of the experiment. The prediction horizon is 20 days. Red points: model prediction; blue points: computational experimental result.

From Fig. 8b to 8d, it is found that compared to the experiment whose update frequency is once per day (Fig. 3), large deviations between simulation predictions and experiment results are observed in the experiment which re-estimates parameter values once every two days. It is then reasonable to explain why this experiment has a much lower hydrogen production, as the accuracy of the model in the experiment is worse. When extending the on-line optimisation prediction horizon, larger errors will be induced if the low accuracy model is used for optimal control estimation. As a result, a longer prediction horizon will eventually lead to a worse optimal control and lower hydrogen production. Hence, in future work it will be necessary to update model parameter values once per day.

3.5 Effect of terminal region and terminal cost

The current work also estimated the optimal hydrogen production when the terminal cost (quadratic objective function presented in Equation 3(b)) and the terminal region constraint that hydrogen production at the end of the prediction horizon must be higher than 100 mL L^{-1} are included in the EMPC. It is found that optimal hydrogen productions are improved by approximate 4% compared to previous implementations in this work when the experiment prediction horizon ranges between 5 days and 20 days, which indicates that both the quadratic terminal cost and the terminal region should be selected in the future. In the current work, the terminal cost takes the form of an unreachable setpoint (2000 mL L^{-1} of total hydrogen production). The use of unreachable setpoints has been theoretically explored for MPC in (Grüne & Pannek 2009). Although there has been no theoretical extension of unreachable setpoints for NMPC or EMPC, for this process it is efficient in practice.

3.6 Comparison of on-line optimisation and off-line optimisation

The off-line optimisation result of the fed-batch computational experiment has also been estimated in the current work. In the off-line optimisation, the initial conditions are kept same as those in the on-line optimisation. The offline optimisation result was obtained by optimising the operating conditions before the commencement of the computational experiment. This method was also implemented in IPOPT in the current study (Wächter & Biegler 2005). It is estimated that the optimal hydrogen production from the off-line optimisation is 658 mL L^{-1}, 13% higher than the on-line optimisation biogas production (582 mL L^{-1}).

It is not surprising to see that off-line optimisation offers a better optimal result compared to on-line optimisation in theory, since on-line optimisation aims to maximise the target production in a prediction horizon instead of the whole process. However, in practice off-line optimisation needs a robust dynamic model characterised by very high accuracy; otherwise large errors will be induced and the optimal control will fail to provide a feasible action to enhance process productivity. This is similar with the case shown in Section 3.4.

Compared to the off-line optimisation, on-line optimisation shows very promising results (88.5% production compared to the off-line) and does not require a highly accurate model to predict the entire performance of an unknown process before its implementation. Therefore, this method is more flexible and accurate than the off-line optimisation method in practice.

Conclusion

In the current study, a novel real-time optimisation strategy which incorporates an economic model predictive control approach and a finite-data window least-squares formulation has been developed to maximise the hydrogen production in a 21-day fed-batch computational experiment. Based on this strategy, a 28.7% increase on hydrogen productivity has been found compared to the previous batch process.

By exploring the impact of essential factors in the on-line optimisation strategy on biogas production, it is concluded that an estimation window length of more than 3 days and a

prediction horizon length of more than 5 days do not significantly affect the on-line optimisation efficiency, and thus the estimation window length and prediction horizon length will be set to these values in future experimental work. It is also concluded that equally spacing controls in the prediction horizon can moderately enhance hydrogen production.

On the contrary, model modification frequency has been found to mainly affect the efficiency of current on-line optimisation methodology, since future process prediction and optimal control decision are strongly dependent on the accuracy of the dynamic model. Therefore, it is essential to re-estimate model parameter values once per day in future work. Furthermore, both the terminal region and the terminal cost selected in the current study will be applied in future work as their implementation is capable of increasing the efficiency of the current optimisation strategy.

In terms of future work, since the real-time optimisation strategy shows a promising biogas production and higher accuracy than the off-line optimisation, real experiments will be further conducted to verify the feasibility and efficiency of the current methodology.

Acknowledgement

E. A. del Rio-Chanona is funded by CONACyT scholarship No. 522530 and from the Secretariat of Public Education and the Mexican government.

References

Adesanya, V.O. et al., 2014. Kinetic modelling of growth and storage molecule production in microalgae under mixotrophic and autotrophic conditions. *Bioresource technology*, 157, pp.293–304. Available at: http://www.ncbi.nlm.nih.gov/pubmed/24576922 [Accessed June 12, 2015].

Alagesan, S. et al., 2013. Model based optimization of high cell density cultivation of nitrogen-fixing cyanobacteria. *Bioresource technology*, 148, pp.228–33. Available at: http://www.ncbi.nlm.nih.gov/pubmed/24047683 [Accessed June 12, 2015].

Anon, 2013. *Model Predictive Contro Lecture Notes, Tech. Rep.*, Available at: http://www.eng.ox.ac.uk/ conmrc/mpc/mpc1-2.pdf.

Arora, N. & Biegler, L.T., 2001. Redescending estimators for data reconciliation and parameter estimation. *Computers and Chemical Engineering*, 25(11-12), pp.1585–1599.

Bandyopadhyay, A. et al., 2010. High rates of photobiological H2 production by a cyanobacterium under aerobic conditions. *Nature communications*, 1, p.139. Available at: http://www.ncbi.nlm.nih.gov/pubmed/21266989 [Accessed June 3, 2014].

Basak, N. & Das, D., 2006. The Prospect of Purple Non-Sulfur (PNS) Photosynthetic Bacteria for Hydrogen Production: The Present State of the Art. *World Journal of Microbiology and Biotechnology*, 23(1), pp.31–42. Available at: http://link.springer.com/10.1007/s11274-006-9190-9.

Basak, N., Jana, A.K. & Das, D., 2014. Optimization of molecular hydrogen production by Rhodobacter sphaeroides O.U.001 in the annular photobioreactor using response surface methodology. *International Journal of Hydrogen Energy*, 39(23), pp.11889–11901. Available at: http://linkinghub.elsevier.com/retrieve/pii/S0360319914014906 [Accessed June 12, 2015].

Bemporad, A. & Morari, M., 1999. Control of systems integrating logic, dynamics, and constraints. *Automatica*, 35(3), pp.407–427.

Biegler, L.T., 1998. Nonlinear Programming. *OR/MS Today*, 25(3), pp.36–45.

Biegler, L.T., 2014. Nonlinear programming strategies for dynamic chemical process optimization. *Theoretical Foundations of Chemical Engineering*, 48(5), pp.541–554. Available at: http://link.springer.com/10.1134/S0040579514050157 [Accessed June 12, 2015].

Boran, E. et al., 2010. Biological hydrogen production by Rhodobacter capsulatus in solar tubular photo bioreactor. *Journal of Cleaner Production*, 18, pp.S29–S35. Available at: http://linkinghub.elsevier.com/retrieve/pii/S0959652610001265 [Accessed January 11, 2015].

Dechatiwongse, P. et al., 2014. Effects of light and temperature on the photoautotrophic growth and photoinhibition of nitrogen-fixing cyanobacterium Cyanothece sp. ATCC 51142. *Algal Research*, 5(0), pp.103–111. Available at:

http://linkinghub.elsevier.com/retrieve/pii/S2211926414000563 [Accessed July 14, 2014].

Dechatiwongse, P., Maitland, G. & Hellgardt, K., 2015. Demonstration of a two-stage aerobic/anaerobic chemostat for the enhanced production of hydrogen and biomass from unicellular nitrogen-fixing cyanobacterium. *Algal Research*, 10, pp.189–201. Available at: http://linkinghub.elsevier.com/retrieve/pii/S2211926415001241 [Accessed June 8, 2015].

Ellis, M. & Christofides, P.D., 2013. Optimal Time-varying Operation of Nonlinear Process Systems with Economic Model Predictive Control. *Industrial & Engineering Chemistry Research*, p.130306145343001. Available at: http://pubs.acs.org/doi/abs/10.1021/ie303537e.

Feng, X. et al., 2010. Mixotrophic and photoheterotrophic metabolism in Cyanothece sp. ATCC 51142 under continuous light. *Microbiology (Reading, England)*, 156(Pt 8), pp.2566–74. Available at: http://www.ncbi.nlm.nih.gov/pubmed/20430816 [Accessed September 10, 2014].

Grüne, L. & Pannek, J., 2011. Nonlinear Model Predictive Control. In *Nonlinear Model Predictive Control*. pp. 67–85. Available at: http://www.springerlink.com/index/10.1007/978-0-85729-501-9.

Grüne, L. & Pannek, J., 2009. Practical NMPC suboptimality estimates along trajectories. *Systems and Control Letters*, 58(3), pp.161–168.

Kameswaran, S. & Biegler, L.T., 2008. Convergence rates for direct transcription of optimal control problems using collocation at Radau points. *Computational Optimization and Applications*, 41(1), pp.81–126.

Kim, J.P. et al., 2010. Repeated production of hydrogen by sulfate re-addition in sulfur deprived culture of Chlamydomonas reinhardtii. *International Journal of Hydrogen Energy*, 35(24), pp.13387–13391. Available at: http://linkinghub.elsevier.com/retrieve/pii/S0360319909019466 [Accessed January 11, 2015].

Lee, C.-M., Hung, G.-J. & Yang, C.-F., 2011. Hydrogen production by Rhodopseudomonas palustris WP 3-5 in a serial photobioreactor fed with hydrogen fermentation effluent. *Bioresource technology*, 102(18), pp.8350–6. Available at: http://www.ncbi.nlm.nih.gov/pubmed/21600763 [Accessed January 12, 2015].

Mailleret, L., Bernard, O. & Steyer, J.-P., 2004. Nonlinear adaptive control for bioreactors with unknown kinetics. *Automatica*, 40(8), pp.1379–1385. Available at: http://linkinghub.elsevier.com/retrieve/pii/S0005109804000792 [Accessed May 14, 2015].

Melis, A. et al., 2000. Sustained photobiological hydrogen gas production upon reversible inactivation of oxygen evolution in the green alga Chlamydomonas reinhardtii. *Plant physiology*, 122(1), pp.127–136. Available at:

http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=58851&tool=pmcentrez&rendertype=abstract.

Min, H. & Sherman, L. a, 2010. Hydrogen production by the unicellular, diazotrophic cyanobacterium Cyanothece sp. strain ATCC 51142 under conditions of continuous light. *Applied and environmental microbiology*, 76(13), pp.4293–4301. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2897434&tool=pmcentrez&rendertype=abstract.

Oh, Y., 2004. Photoproduction of hydrogen from acetate by a chemoheterotrophic bacterium Rhodopseudomonas palustris P4. *International Journal of Hydrogen Energy*, 29(11), pp.1115–1121. Available at: http://linkinghub.elsevier.com/retrieve/pii/S0360319903003306.

Del R ó-Chanona, E.A. et al., 2015. Optimal Operation Strategy for Biohydrogen Production. *Industrial & Engineering Chemistry Research*, p.150602155211002. Available at: http://pubs.acs.org/doi/abs/10.1021/acs.iecr.5b00612 [Accessed June 12, 2015].

Tamburic, B. et al., 2012a. A novel nutrient control method to deprive green algae of sulphur and initiate spontaneous hydrogen production. *International Journal of Hydrogen Energy*, 37(11), pp.8988–9001. Available at: http://linkinghub.elsevier.com/retrieve/pii/S0360319912003655 [Accessed March 24, 2014].

Tamburic, B. et al., 2011. Design of a novel flat-plate photobioreactor system for green algal hydrogen production. *International Journal of Hydrogen Energy*, 36(11), pp.6578–6591. Available at: http://linkinghub.elsevier.com/retrieve/pii/S0360319911004459 [Accessed November 12, 2014].

Tamburic, B. et al., 2012b. Effect of the Light Regime and Phototrophic Conditions on Growth of the H2-producing Green Alga Chlamydomonas Reinhardtii. *Energy Procedia*, 29(0), pp.710–719. Available at: http://linkinghub.elsevier.com/retrieve/pii/S1876610212015032 [Accessed November 13, 2014].

Tamburic, B. et al., 2013. Process and reactor design for biophotolytic hydrogen production. *Physical chemistry chemical physics : PCCP*, 15(26), pp.10783–94. Available at: http://www.ncbi.nlm.nih.gov/pubmed/23689756 [Accessed November 12, 2014].

Vijayaraghavan, K., Karthik, R. & Kamala Nalini, S.P., 2009. Hydrogen production by Chlamydomonas reinhardtii under light driven sulfur deprived condition. *International Journal of Hydrogen Energy*, 34(19), pp.7964–7970. Available at: http://linkinghub.elsevier.com/retrieve/pii/S0360319909012567 [Accessed January 11, 2015].

Wächter, A. & Biegler, L.T., 2005. *On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming*, Available at: http://link.springer.com/10.1007/s10107-004-0559-y.

Wang, Y.-Z. et al., 2013. Bioconversion characteristics of Rhodopseudomonas palustris CQK 01 entrapped in a photobioreactor for hydrogen production. *Bioresource technology*, 135, pp.331–8. Available at: http://www.ncbi.nlm.nih.gov/pubmed/23127839 [Accessed June 12, 2015].

Zavala, V.M. & Biegler, L.T., 2006. Large-scale parameter estimation in low-density polyethylene tubular reactors. *Industrial and Engineering Chemistry Research*, 45(23), pp.7867–7881.

Zhang, D., Del-Rio Chanona, E.A., et al., 2015. Analysis of green algal growth via dynamic model simulation and process optimisation. *Biotechnology and bioengineering*. Available at: http://www.ncbi.nlm.nih.gov/pubmed/25855209 [Accessed May 3, 2015].

Zhang, D., Dechatiwongse, P., et al., 2015. Analysis of the Cyanobacterial Hydrogen Photo-Production Process via Model Identification and Process Simulation. *Chemical Engineering Science*, 128, pp.130–146. Available at: http://linkinghub.elsevier.com/retrieve/pii/S0009250915000883 [Accessed February 16, 2015].

Zhang, D. et al., 2015. Bioprocess modelling of biohydrogen production by Rhodopseudomonas palustris: Model development and effects of operating conditions on hydrogen yield and glycerol conversion efficiency. *Chemical Engineering Science*, In Press. Available at: http://linkinghub.elsevier.com/retrieve/pii/S0009250915001815 [Accessed March 24, 2015].

Zhang, D., Dechatiwongse, P., et al., 2015. Dynamic modelling of high biomass density cultivation and biohydrogen production in different scales of flat plate photobioreactors. *Biotechnology and bioengineering*. Available at: http://www.ncbi.nlm.nih.gov/pubmed/26041472 [Accessed June 12, 2015].

Zhang, Z. & Chen, J., 2015. Correntropy based data reconciliation and gross error detection and identification for nonlinear dynamic processes. *Computers & Chemical Engineering*, 75, pp.120–134. Available at: http://linkinghub.elsevier.com/retrieve/pii/S009813541500006X.